(72) Inventors:
• Srikanth, Ayikudy K.
Reading, Massachusetts 01867 (US)
• Meehan, Tom
New York, New York 10002 (US)
• Mujeeb, Harmayon
Chelmsford, Massachusetts 01824 (US)

(74) Representative: Dearling, Bruce Clive et al
Hepworth Lawrence Bryer & Bizley,
Merlin House,
Falconry Court,
Bakers Lane
Epping, Essex CM16 5DQ (GB)

(54) **Method and apparatus providing for router redundancy of non internet protocols using the virtual router redundancy protocol**

(57)   One or more nodes configured with data communication protocol suites other than TCP/IP, and for which a default router election protocol such as VRRP is not available, forwards non TCP/IP traffic destined to nodes on other network to the VRRP master virtual router. A non TCP/IP based node, given a statically configured Network layer address for a next hop router, resolves the Network layer address to VRRP based MAC layer address for the next hop router, if the next hop router has VRRP enabled and is the master virtual router for TCP/IP traffic. The non TCP/IP node caches the VRRP based MAC address for use in forwarding non TCP/IP traffic to an available next hop router. Thereafter, the non TCP/IP node forwards traffic destined to nodes on other networks to the VRRP master virtual router. If the master virtual router becomes unavailable to forward TCP/IP based traffic, it transitions to become the new backup virtual router, while the backup virtual router transitions to become the new master virtual router responsible for forwarding TCP/IP based traffic. The new master virtual router shares the same VRRP based MAC address. Therefore, the non TCP/IP based node, having cached the VRRP based MAC address of the next hop router, forwards traffic destined to nodes on other networks to the new master virtual router. This removes the need for intervention or reconfiguration of the non TCP/IP based node for continued routing of non TCP/IP traffic transmitted therefrom in the event that the router identified therein by the statically configured Network layer address for a next hop router becomes unavailable.
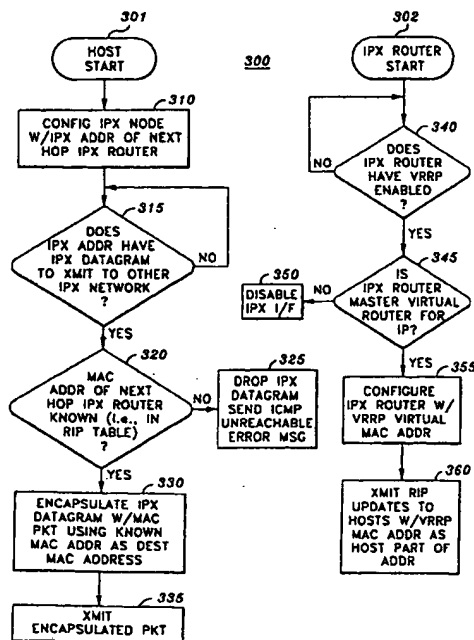
FIG. 3

EP 1 011 231 A2

Description

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention is related to data communications. In particular, the present invention is related to providing router redundancy in a statically configured routing environment for multiple protocols using the virtual router redundancy protocol (VRRP).

Description of the Related Art

[0002] The Transport Control Protocol/Internet Protocol (TCP/IP) suite of data communication protocols is used in many of today's internet networks (internets). A TCP/IP-based internet provides a data packet switching system for communication between nodes (e.g., end-user workstations, servers, network devices, etc.) connected to the internet. With reference to Figure 1, International Standards Organization (ISO) Open Systems Interconnection (OSI) Network-layer devices 105, 110, and 140, known as routers or switches, select a path and forward, i.e., route, IP datagrams between nodes connected to the internet 100. For example, internet 100 includes local area networks (LANs) 101 and 151, and wide area network (WAN) 102 interconnected by routers 105, 110 and 140. The routers route IP datagrams, for example, between nodes 115, 120, 125 and 130 attached to LAN 101 and nodes 145 and 150 attached to LAN 151.

[0003] As can be seen from Figure 1, routers 105 and 110 provide multiple paths for transmitting IP datagrams from nodes on LAN 101 to nodes on other IP networks in the internet. To prevent generating and forwarding duplicate IP datagrams over the internet, each of the nodes on LAN 101 transmits a unicast IP datagram to only one of routers 105 and 110 as a next hop, or first hop, router. The next hop router forwards the datagram to a node on another IP network in the internet that is reachable via the router. As is well known in the art, a next hop router can be statically configured at each node as the default router (also referred to as the default gateway) towards another IP network. However, a static default router configuration provides a single point of failure in the event the default router configuration provides a single point of failure in the event the default router becomes unavailable. To overcome this problem, next hop routers can be dynamically configured at each node, using a dynamic routing protocol such as the well known Routing Information Protocol (RIP) or Open Shortest Path First (OSPF) dynamic routing protocols. However, the reliability provided by a dynamic routing protocol is at the expense of node and router processing overhead, network overhead, interoperability problems, etc.

[0004] The single point of failure inherent in a static next hop router configuration can be overcome through the use of the Virtual Router Redundancy protocol (VRRP), VRRP, as set forth in the Internet Society's Request For Comments 2338, April, 1998, is an election protocol that assigns responsibility to a master virtual router, wherein the master virtual router is one of two or more VRRP based routers attached to a LAN. VRRP provides dynamic fail-over in forwarding responsibility if the master virtual router, selected as one of the two or more VRRP routers on the LAN, becomes unavailable. In essence, and as explained in detail in RFC 2338, VRRP provides a redundant, relatively more reliable default path for transmission of IP datagrams destined to nodes on other IP networks.

[0005] For a better appreciation and understanding of the present invention, a brief review of the VRRP protocol follows. In the network illustrated in Fig. 2, routers 105 and 110 operate according to the VRRP. VRRP is based on the concept of a virtual router, which is an abstract object that operates as a default router for nodes attached to the LAN, and for which a single well known IEEE 802.3 MAC (Media Access Control) address is assigned. In network 100, two virtual routers are configured: virtual router 1, and virtual router 2. In general, the scope of a virtual router is restricted to a single LAN, and each virtual router comprises a master and one or more backup routers. For example, router 105 is the master virtual router and router 110 is the backup virtual router for virtual router 1. The master and backup virtual routers share the same virtual router identifier (VRID = 1), same IP address (IP = A), and the same VRRP-based MAC address (00-00-5E-00-01-{VRID}(h), e.g., 00-00-5E-00-01-01(h)). Conversely, router 110 is the master virtual router and router 105 is the backup virtual router for virtual router 2, which has a VRID = 2, IP address = B, and VRRP-based MAC address of 00-00-5E-00-01-02(h). In this example, nodes 115 and 120 are statically configured with a default next hop router IP address of "A", while nodes 125 and 130 are statically configured with a default next hop router IP address of "B". (Splitting the nodes between redundant routers in this manner provides load balancing and other advantages well known in the art.)

[0006] In accordance with VRRP, the master virtual router functions as the forwarding router for the IP address associated with the virtual router. For example, router 105 is the master virtual router for virtual router 1, and forwards IP datagrams for the nodes having the IP address of virtual router 1 (IP address "A") statically configured as the default next hop router. The master virtual router periodically transmits advertisements, formatted as IP multicasts, to the backup virtual router(s) on the local network to indicate to the backup(s) that it is still functioning as the master virtual router. (Use of IP multicast allows VRRP to be implemented in multiaccess LANs such as Ethernet, Fast Ethernet, Gigabit Ethernet, FDDI, Token Ring, ATM LAN emulation (ATM LANE), etc.) If master virtual router 1 fails, e.g., advertisements cease, the master-down_timer expires or a

shutdown event is received at the backup virtual router 1, backup virtual router 1 takes over as the new master virtual router 1, providing routing capability for nodes 115 and 120. Since both routers maintain the same IP address ("IP A"), and share the same VRRP based media access control (MAC) address, no reconfiguration of the static default next hop router IP address is required at each of the nodes that transmit IP datagrams to virtual router 1, destined for nodes on other IP networks. Likewise, if master virtual router 2 fails, backup virtual router 2 provides routing for nodes 125 and 130.

[0007] While the VRRP based network of Figure 2 provides router redundancy and fail-over protection for nodes having statically configured default next hop router IP addresses, no such mechanism presently exists for protocol stacks other than TCP/IP. However, many of today's internets are a connected amalgamation of heterogeneous, originally isolated, data communication networks, wherein the routers often times are required to support multiprotocol routing capabilities, including data communication protocols for which a default router election protocol such as VRRP is not available. For example, in the internet depicted in Fig. 2, nodes 115 and 125 communicate with each other as well as nodes reachable via router 105 using both IP and the well known Network layer protocol Internet Protocol eXchange (IPX), yet VRRP does not provide routing redundancy support for IPX. What is needed is a system in which nodes utilizing data communication protocols in addition to the Internet Protocol (IP) can obtain the benefits of VRRP without implementing a separate default router election protocol.

BRIEF SUMMARY OF THE INVENTION

[0008] According to the present invention, nodes configured with data communication protocol suites other than TCP/IP, for which a default router election protocol such as VRRP is not available, nevertheless forward traffic destined to nodes on other network to the VRRP master virtual router. A non TCP/IP based node, given a statically configured Network layer address for a next hop router, resolves the Network layer address to a VRRP based MAC layer address for the next hop router, if the next hop router has VRRP enabled and is the master virtual router. The non TCP/IP node caches the VRRP based MAC address for use in forwarding non TCP/IP traffic to the appropriate next hop router. Thereafter, the non TCP/IP node forwards traffic destined to nodes on other networks to the VRRP master virtual router. If the master virtual router becomes unavailable to forward TCP/IP based traffic, it transitions to become the new backup virtual router, while the backup virtual router transitions to become the new master virtual router responsible for forwarding TCP/IP based traffic. The non TCP/IP based node, having cached the VRRP based MAC address of the next hop router, forwards traffic destined to nodes on other networks to the new master vir-

tual router, which shares the same VRRP based MAC address as the former master virtual router.

BRIEF SUMMARY OF THE SEVERAL VIEWS OF THE DRAWINGS

[0009] The present invention is illustrated by way of example and not limitation in the following figures, in which:
[0010] Figure 1 is a diagram of a data communications internetwork.
[0011] Figure 2 is a diagram of a data communications internetwork that supports the virtual router redundancy protocol.
[0012] Figure 3 is a flow chart of an embodiment of a method of the present invention.
[0013] Figure 4 is a flow chart of an embodiment of a method of the present invention.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

[0014] Described is a method and apparatus providing for nodes utilizing data communication protocols, e.g., Internet Protocol eXchange (IPX), or Internet Gateway Management Protocol (IGMP), in addition to the Internet Protocol (IP) can obtain the benefits of VRRP without implementing a separate default router election protocol.
[0015] In the following description, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, to one of ordinary skill in the art that the present invention may be practised without these specific details. In other instances, well-known architectures, steps, and techniques have not been shown to avoid unnecessarily obscuring the present invention. For example, specific details are not provided as to whether the method is implemented in a switch as a software routine, hardware circuit, firmware, or a combination thereof.
[0016] In alternative embodiments, the present invention may be applicable to implementations of the invention in integrated circuits or chip sets, wireless implementations, switching systems products and transmission systems products. For purposes of this application, the terms switching systems products shall be taken to mean private branch exchanges (PBXs), central office switching systems that interconnect subscribers, toll/tandem switching systems for interconnecting trunks between switching centers, and broadband core switches found at the center of a service provider's network that may be fed by broadband edge switches or access multiplexors, and associated signalling, and support systems and services. The term transmission systems products shall be taken to mean products used by service providers to provide interconnection between their subscribers and their networks such as loop systems,

and which provide multiplexing, aggregation and transport between a service provider's switching systems across the wide area, and associated signalling and support systems and services.

[0017]   In an internetwork operating in accordance with VRRP as depicted, for example, in Fig. 2, the present invention provides for route redundancy and fail-over protection for nodes having statically configured default next hop router addresses for protocol stacks other than TCP/IP. With reference to Fig. 3, an embodiment of a method of the present invention is set forth, using the IPX protocol as an example. However, it is appreciated that the present invention is applicable in the same manner with many other data communication protocols, e.g., IGMP relay, etc. The method is set forth in flow charts describing the process from the perspective of a node, beginning at 301, and from the perspective of the next hop IPX router, beginning at 302.

[0018]   In the internet depicted in Fig. 2, nodes may communicate with each other via multiple Network layer protocols, e.g., IP and IPX. In particular, nodes 115 and 125 may communicate with nodes in or reachable beyond WAN 102 via routers 105 and 110, using either IP or IPX at the Network layer. At step 310, node 115 is statically configured with an IP address of "A" for a next hop IP router, and an IPX address of "X" for a next hop IPX router. As is well known in the art, when a source node sends a datagram to a destination node, the Network-layer destination address is that of the destination node. If, however, the Network layer destination address indicates that the node is on a different network (and, therefore, reachable only by a Network layer router), then the MAC layer destination address is not that of the node, but rather that of the next hop router. Initially, e. g., on power up, the sending, or source, node knows the statically configured Network layer address of the next hop router, but not the MAC address of the next hop router. Both IPX and IP have well known mechanisms for resolving Network layer addresses to MAC addresses, and invoke such mechanisms, either upon boot up, or when transmitting a first datagram to a destination node in another network. Thus, at step 315, node 11 5, for example, waits until it has an IPX datagram to transmit to another node in a different IPX network before it checks, e.g., an internal table or cache such as a Routing Information Protocol (RIP) table, at step 320, for a MAC address associated with the IPX address (X) of the next hop IPX router.

[0019]   If the MAC address of the next hop IPX router is in the node's cache, at step 330, the IPX datagram is encapsulated in a MAC packet, wherein the destination MAC address is the MAC address of the next hop IPX router. The MAC packet is then transmitted at step 335 to the next hop IPX router. (The next IPX hop router, of course, strips the MAC packet to retrieve the IPX datagram, looks in its routing tables for the subsequent next hop IPX router to which to forward the IPX datagram based on the destination IPX address in the datagram,

encapsulates the IPX datagram in a MAC packet, wherein the destination MAC address is the MAC address of the subsequent next hop IPX router, or the destination node itself, if the destination node is directly reachable, i.e., attached to a network to which the next hop IPX router is attached).

[0020]   If at step 320 the MAC address of the next hop IPX router is not in the source node's cache, the source node drops the IPX datagram and sends an Internet Control Message Protocol (ICMP) unreachable error message up the IPX protocol stack executing at the source node, as shown at step 325.

[0021]   Recall from the above discussion that router 105 is a multiprotocol router, providing routing functionality for multiple protocols, e.g., both the IPX and IP protocols. Moreover, the router, in conjunction with router 110, has VRRP enabled to provide dynamic fail-over capabilities for IP in the event the router 105 becomes unavailable to route IP traffic. Thus, in accordance with the present invention, rather than sending out an IPX RIP update packet with the actual MAC address of its appropriate port, router 105 first determines at step 340 whether VRRP is enabled. If router 105 has VRRP enabled, it further determines at step 345 whether it is the master virtual router for IP traffic. If router 105 has VRRP enabled, and is the master virtual router for IP traffic, then at step 355, router 105 configures itself with the appropriate VRRP based MAC address, overriding the physical MAC address assigned to the port at which VRRP is enabled. At 360, router 105 then transmits its VRRP-based MAC address (e.g., 00-00-5E-00-01-01(h)) in its IPX RIP update packets. If in router 105 VRRP is enabled, or if VRRP is enabled but the router is the backup virtual router for IP traffic, then at step 350, router 105 disables IPX routing at the port for which VRRP is enabled. Moreover, if VRRP is disabled, the IPX interface comes back up and the router transmits its actual MAC address in its IPX RIP update packets.

[0022]   In any case, node 115 caches the MAC address received from router 105, for example, in a RIP table or like data structure. Then, as discussed above, once the MAC address of the next hop IPX router is cached, whether that be the actual MAC address or the VRRP based MAC address, at step 330, the IPX datagram is encapsulated in a MAC packet, wherein the destination MAC address is the MAC address of the next hop IPX router stored in cache. The packet is then transmitted at step 335 to the next hop IPX router.

[0023]   Thereafter, each time node 115 transmits IPX datagrams to another node in another IPX network, it routes the IPX datagrams to the next hop IPX router based on the MAC address stored in its cache. If the cache entry times out or expires, the process is repeated. Importantly, if router 105 becomes unavailable to route IP traffic, VRRP transitions between routers 105 and 110 so that router 110 becomes the new master virtual router and router 105 becomes the new backup virtual router for IP traffic. Since the new master virtual

router uses the same VRRP based MAC address as the original master virtual router, the MAC address already stored in cache at node 115 is the VRRP based MAC address of router 110 as well. Thus, after the transition, all IPX traffic sourced by node 115 and destined for a node in another IPX network is automatically routed to that other IPX network via router 110, since it has the same VRRP based MAC address as router 105.

[0024] It is appreciated then, that depending on the IPX address of the next hop IPX router statically configured at each node in a shared network, and depending on the VRRP status of the next hop IP router when the node obtains a RIP update from the next hop IP router, it is quite possible that some IPX based nodes will forward all IPX traffic destined to another IPX network to the same next hop IPX router (if the router's actual MAC address is cached), while other IPX based nodes will forward all IPX traffic destined to another IPX network to a different next hop router (if that router's actual MAC address is cached), while still other IPX based nodes will forward IPX traffic destined to another IPX network to the one of the VRRP based routers that is the current master virtual IP router (if the router's VRRP-based MAC address is cached).

[0025] In Figure 4, an embodiment 400 of the process of the present invention as it relates to another protocol, the Internet Group Management Protocol (IGMP), is provided. Router 105, in conjunction with router 110, is configured to enable VRRP to provide dynamic fail-over capabilities for IP in the event the router 105 becomes unavailable to route IP traffic. Thus, in accordance with the present invention, router 105 first determines at step 410 whether VRRP is enabled. If router 105 has VRRP enabled, it further determines at step 420 whether it is the master virtual router for IP traffic. If router 105 has VRRP enabled, and is the master virtual router for IP traffic, then at step 440, router 105 enables IGMP Relay at the port where VRRP is enabled and in the master state. If VRRP is enabled but the router is the backup virtual router for IP traffic, then at step 430, router 105 disables IGMP relay at the port for which VRRP is enabled.

[0026] Whenever node 115 multicasts an IGMP packet, router 105 would be in the master state for the virtual router, and hence, the IGMP Relay interface would be enabled on that node to forward IGMP traffic. Router 110 would be in backup state and so the IGMP interface on that router would be disabled. When router 105 becomes unavailable, the IGMP interface on router 105 is disabled. Additionally, router 110 takes the role of the master, and the appropriate IGMP interface is enabled on router 110 to forward IGMP traffic received at the IGMP interface.

[0027] Embodiments of the invention may be represented as a software product stored on a machine-readable medium (also referred to as a computer-readable medium or a processor-readable medium). The machine-readable medium may be any type of magnetic, optical, or electrical storage medium including a diskette, CD-ROM, memory device (volatile or non-volatile), or similar storage mechanism. The machine-readable medium may contain various sets of instructions, code sequences, configuration information, or other data. For example, the procedures described above for synchronizing network address translation tables can be stored on the machine-readable medium. Those of ordinary skill in the art will appreciate that other instructions and operations necessary to implement the described invention may also be stored on the machine-readable medium.

Claims

1. A method for changing in a network providing for a statically configured next hop router for a first protocol, a default next hop router for a second protocol, comprising:

   a) initialising a router connected to the network as the master next hop router for the first protocol;
   b) initialising a redundant router connected to the network as the backup next hop router for the first protocol;
   c) initialising the router as the default next hop router for the second protocol;
   d) transitioning to the redundant router as the default next hop router for the second protocol upon the occurrence of an event causing a transition to the redundant router as a new master next hop router and a transition to the master next hop router as a new backup next hop router for the first protocol.

2. The method of claim 1, wherein the occurrence of the event causing a transition to the redundant router as a new master next hop router and a transition to the master next hop router as a new backup next hop router for the first protocol comprises one of:

   a failure of the master next hop router to advertise itself as such within an appropriate time interval to the backup next hop router;
   a pre-emption of the master next hop router by the backup next hop router; and
   a power failure of the master next hop router.

3. A computer program product stored on a computer usable medium, for execution on a router connected to a network, comprising:

   means for initialising a router connected to the network as the master next hop router for the first protocol;
   means for initialising a redundant router con-

nected to the network as the backup next hop
router for the first protocol;
means for initialising the router as the default
next hop router for the second protocol;
means for transitioning to the redundant router
as the default next hop router for the second
protocol, if an event causing a transition to the
redundant router as a new master next hop
router and a transition to the master next hop
router as a new backup next hop router for the
first protocol.

4. A network having a router coupled to the network
as a master next hop router for a first protocol and
a redundant router coupled to the network as a
backup next hop router for the first protocol, the net-
work configured to provide a statically configured
next hop router for the first protocol and configura-
ble to change a default next hop router for a second
protocol, the network comprising:

means for initialising the master next hop router
for the first protocol;
means for initialising the redundant router as
the back up next hop router for the first protocol;
means for initialising the router as the default
next hop router for the second protocol; and
means for transitioning to the redundant router
as the default next hop router for the second
protocol, said means for transitioning respon-
sive to an occurrence of an event that causes
a transition to the redundant router as a new
master next hop router and a transition to the
master next hop router as a new back up next
hop router for the first protocol.

5. The method of claim 1 or the network of claim 4,
wherein the first protocol is an Internet Protocol (IP).

6. The method of claim 1 or 5 or the network of claim
4 or 5, wherein the second protocol is one of:

an Internet Protocol eXchange (IPX) protocol;
and
an Internet Gateway Management Protocol
(IGMP).

7. In a network providing for a statically configured
next hop router for a first protocol, a method for
changing a default next hop router for a second pro-
tocol, comprising:

a) initialising a router connected to the network
as the master next hop router for the first pro-
tocol;
b) initialising a redundant router connected to
the network as the backup next hop router for
the first protocol;

c) initialising the router as the default next hop
router for the second protocol;
d) if an event occurs causing a transition to the
redundant router as a new master next hop
router and a transition to the master next hop
router as a new backup next hop router for the
first protocol, then transitioning to the redun-
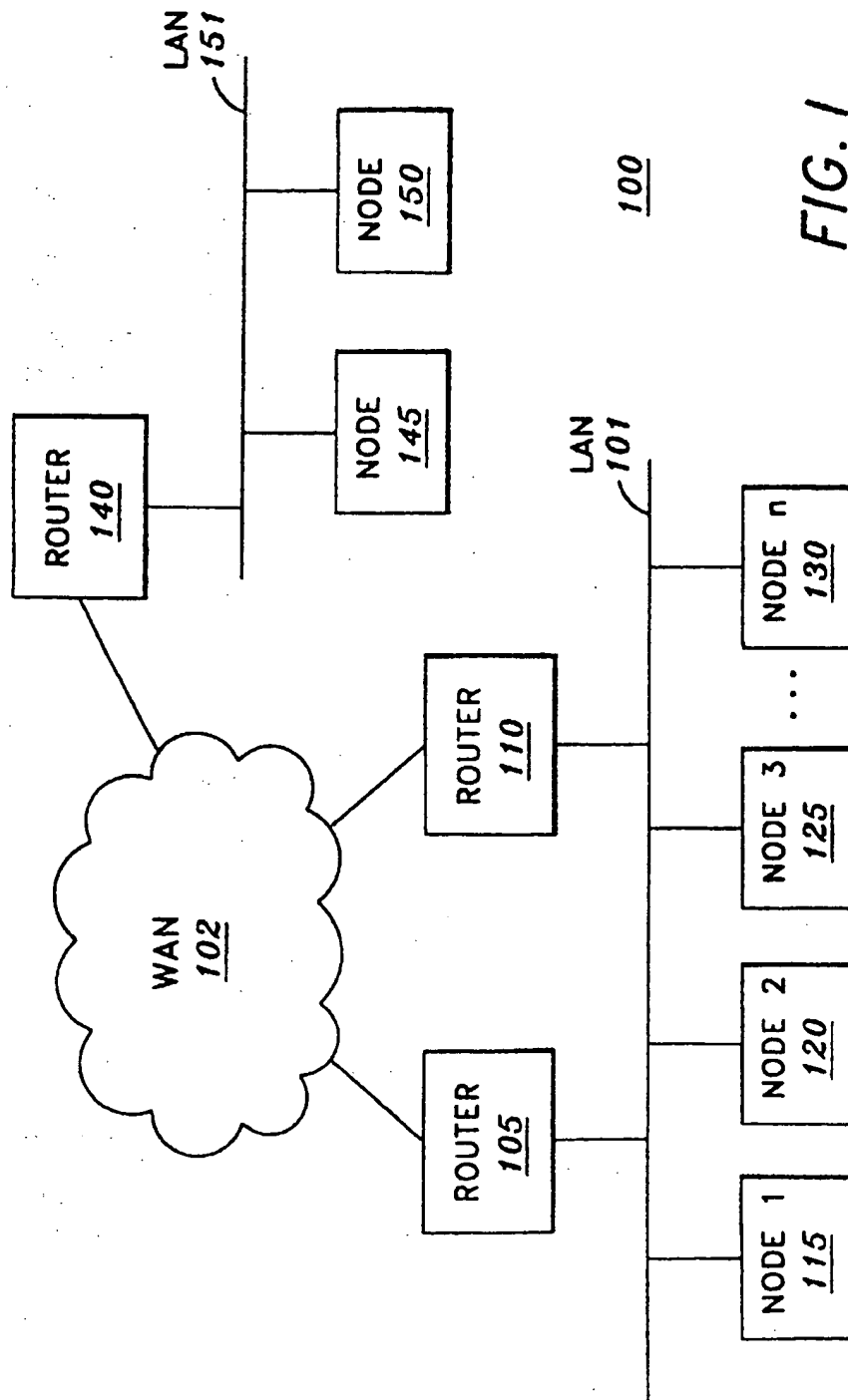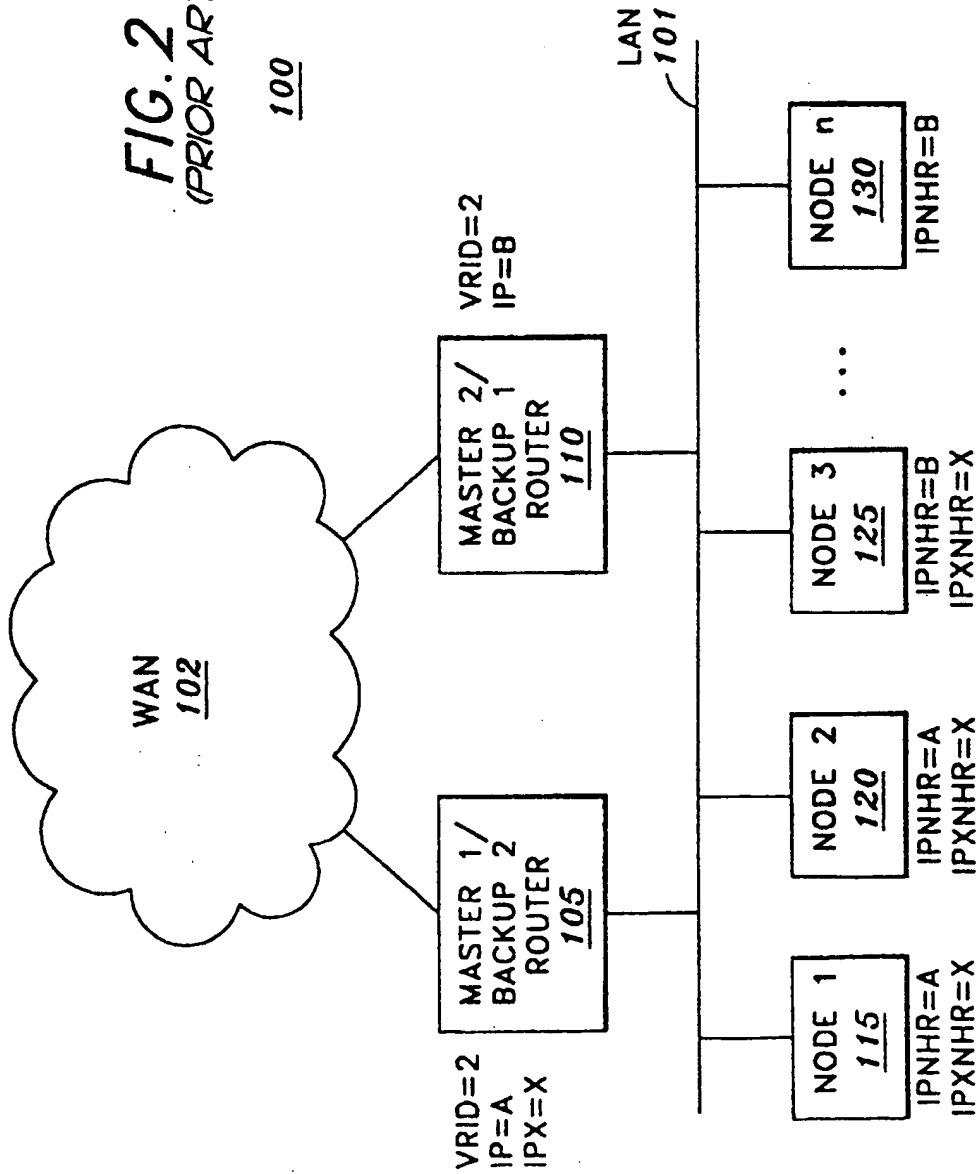dant router as the default next hop router for
the second protocol.

FIG. 1
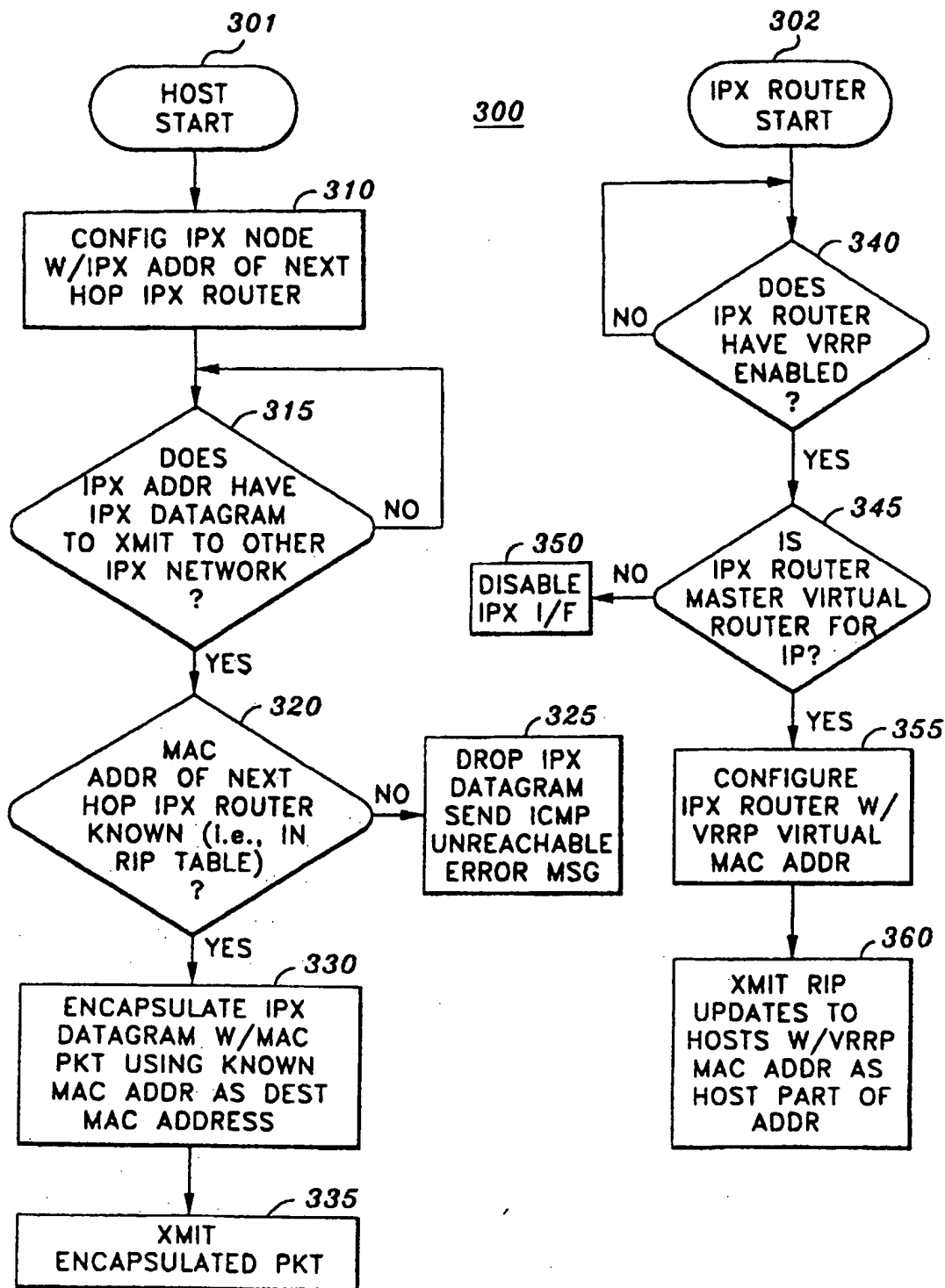(PRIOR ART)

FIG. 2
(PRIOR ART)

100

300

301

HOST
START

310
CONFIG IPX NODE
W/IPX ADDR OF NEXT
HOP IPX ROUTER

315
DOES
IPX ADDR HAVE
IPX DATAGRAM
TO XMIT TO OTHER
IPX NETWORK
?
— NO

YES

320
MAC
ADDR OF NEXT
HOP IPX ROUTER
KNOWN (I.e., IN
RIP TABLE)
?
— NO

325
DROP IPX
DATAGRAM
SEND ICMP
UNREACHABLE
ERROR MSG

YES

330
ENCAPSULATE IPX
DATAGRAM W/MAC
PKT USING KNOWN
MAC ADDR AS DEST
MAC ADDRESS

335
XMIT
ENCAPSULATED PKT

302

IPX ROUTER
START

340
DOES
IPX ROUTER
HAVE VRRP
ENABLED
?
— NO

YES

345
IS
IPX ROUTER
MASTER VIRTUAL
ROUTER FOR
IP?
— NO

350
DISABLE
IPX I/F

YES

355
CONFIGURE
IPX ROUTER W/
VRRP VIRTUAL
MAC ADDR

360
XMIT RIP
UPDATES TO
HOSTS W/VRRP
MAC ADDR AS
HOST PART OF
ADDR

FIG. 3

401

START

410

DOES
IGMP RELAY
BASED ROUTER
HAVE VRRP
ENABLED
?

400

YES

420

IS
IGMP RELAY
BASED ROUTER
MASTER VIRTUAL
ROUTER FOR
IP?

NO

430

DISABLE IGMP
RELAY I/F

YES

440

ENABLE IGMP
RELAY I/F

FIG. 4